IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Appellant	:	Nicholas Lee	Confirmation No.: 8624
Appl. No.	:	09/650,173	
Filed	:	August 29, 2000	
For	:	VOICE INTERFACE AND METHODS FOR IMPROVING RECOGNITION ACCURACY OF VOICE SEARCH QUERIES	
Examiner	:	Opsasnick	

THIRD APPEAL BRIEF

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

This Third Appeal Brief is being filed to address the modified basis for rejection set forth in the Office Action mailed on February 3, 2006 (the "Current Office Action"), which was issued in response to Appellant's Supplemental Appeal Brief filed on September 10, 2004.

The Current Office Action sets forth the same bases for the rejection of Claims 1-55 as the prior Office Action (mailed on August 11, 2004), and also includes a new art-based rejection of Claim 50. The Current Office Action also addresses certain issues raised in Appellant's Supplemental Appeal Brief.

I. REAL PARTY IN INTEREST

The real party in interest in the present application is A9.com, Inc.

II. RELATED APPEALS AND INTERFERENCES

No related appeals, interferences or judicial proceedings are currently pending.

III. STATUS OF CLAIMS

Claims 1-55, which are attached hereto as an appendix, are currently pending in the application and are the subject of this appeal.

IV. STATUS OF AMENDMENTS

No amendments have been made since Appellant filed his original appeal.

V. SUMMARY OF CLAIMED SUBJECT MATTER

The present application includes six independent claims. Each independent claim is briefly summarized and then reproduced below, together with citations to corresponding portions of the specification and drawings as required by 37 C.F.R. § 41.37(c)(1)(v). These citations are provided to illustrate specific examples and embodiments of the recited claim language, and are not intended to limit the claims.

Claim 1

Claim 1 is directed to a process for accurately capturing a search query intended by a user, without requiring the user to type in or utter all of the characters of the search query. The recited process is well suited for conducting searches via a telephone, although the claim is not so limited. The process involves having the user utter (speak) the search query, and also having the user submit a subset of the characters of the search query (e.g., the first three of four characters). The subset of characters may, for example, be submitted via telephone keypad entries and/or voice. To interpret the utterance of the search query (referred to as a "voice query"), the items that match the submitted set of characters are identified, and are used to dynamically generate a voice recognition grammar (referred to as a "dynamic grammar") that specifies valid utterances. This dynamic grammar is then used to interpret the utterance of the search query. Because the dynamic grammar is generated using the items that match the submitted characters, as opposed, e.g., to being generated from all items in the searchable domain, it enables the utterance to be recognized with a high degree of accuracy. See page 2, lines 13-23.

More specifically, Claim 1 is directed to a method for improving voice recognition accuracy when a user submits a search query by voice to search a domain of items, the method comprising:

- prompting a user to submit a set of characters of a voice query for searching the domain of items, and receiving the set of characters from the user, wherein the voice query is an utterance by the user of a search query, and the set of characters defines a portion of the search query (see, e.g., page 1, line 29 to page 2, line 12; Fig. 1, block 22; page 5, lines 1-18; page 6, lines 10-23; Fig. 3, block 80; and page 8, lines 12-30);
- in response to receiving the set of characters from the user, identifying a subset of items in the domain that correspond to the set of characters (see, e.g., page 2, lines 13-23; Fig. 1, block 24; page 5, lines 19-24; Fig. 3, block 84; and page 9, lines 2-7);
- generating a dynamic grammar based at least in part on the subset of items, said grammar specifying valid utterances for interpreting the voice query (see, e.g., page 2, lines 13-23; Fig. 1, block 26; page 5, line 25 to page 6, line 9; Fig. 3, block 86; and page 9, lines 5-15);
- prompting the user to submit the voice query, and receiving the voice query from the user (see, e.g., page 2, line 12; Fig. 1, block 28; page 6, lines 24-26; Fig. 3, block 90; and page 9, lines 16-20);
- interpreting the voice query using the dynamic grammar (see, e.g., page 2, lines 13-20; Fig. 1, block 28; page 6, lines 26-30; Fig. 3, block 90; and page 9, lines 18-20).

Claim 15

Claim 15 differs from Claim 1 primarily in that it does not require the voice recognition grammar to be generated dynamically. For example, Claim 15 would also cover a process in which the voice recognition grammar corresponding to a particular set of characters is pre-generated, and is retrieved from memory in response to the user's input.

More specifically, Claim 15 is directed to a method for improving voice recognition accuracy when a user submits a query by voice to search a domain of items, the method comprising:

- receiving a set of characters entered by a user, the set of characters representing a portion of a query (see, e.g., page 1, line 29 to page 2, line 12; Fig. 1, block 22; page 5, lines 1-18; page 6, lines 10-23; Fig. 3, block 80; and page 8, lines 12-30);
- in response to receiving the set of characters, selecting a grammar which is derived at least in-part from text extracted from a subset of items that correspond to the set of characters entered by the user (see, e.g., page 2, lines 13-23; Fig. 1, blocks 24, 26 and 32; page 5, lines 19-26; page 7, lines1-10; Fig. 3, blocks 82-88; and page 9, lines 2-18) and
- providing the grammar to a voice recognition system for use in interpreting the query as entered by the user by voice (see, e.g., Fig. 2, and page 9, lines 16-20);
- whereby the user's entry of a subset of characters of the query, together with the user's utterance of the full query, are used in combination to capture the query (see, e.g., page 6, lines 24-30 and page 8, lines 12 to page 9, line 20).

Claim 24

Claim 24 is directed to a system that enables a user to submit a search query by uttering the search query, and by also entering a subset of the characters of the search query. The utterance of the search query is again referred to as a "voice query."

More specifically, Claim 24 is directed to a system for conducting searches by voice, the system comprising:

- a database of items (see, e.g., Figure 2, database 62, and page 8, lines 1-4);
- a query server which searches the database of items according to voice queries from users, the query server coupled to a voice recognition system which interprets the voice queries according to grammars (see, e.g., Fig. 2, blocks 50 and 60, page 7, line 23 to page 8, line 15; and Fig. 3);
- a first code module which causes a user to be prompted to enter a set of characters of a query such that the user may partially specify the query (see, e.g., Fig. 1, block 22; Fig. 3, block 80 and page 8, lines 12-20); and
- a second code module which causes the user to be prompted to utter the query (see, e.g., Fig. 1, block 28; Fig. 3, block 90; and page 9, lines 16-18);

- wherein the query server is programmed to use the set of characters to select a grammar for use by the voice recognition system to interpret the query as uttered by the user (see, e.g., page 9, lines 16-23).

Claim 33

Claim 33 is directed to a method by which a user can conduct a search via a voice query, and then refine the search by adding a query term to the query. (Note that Claim 33 is the only independent claim that describes such a search query refinement process.) To interpret the user's utterance of this additional query term, a grammar is generated, at least in-part, by extracting text from the set of search result items resulting from the voice query. This process enables the user's utterance of the additional query term to be recognized with a high degree of accuracy.

More specifically, Claim 33 recites a method of assisting users in locating items in a database using voice queries, the method comprising:

- receiving a voice query from a user, and identifying a set of search result items that are responsive to the voice query (see, e.g., Fig. 1, blocks 28 and 30; Fig. 3, blocks 90 and 92; and page 9, lines 24-30);
- providing the user an option to refine the query by adding an additional query term (see, e.g., Fig. 4, block 94, and page 9, line 27 to page 10, line 2);
- generating a grammar at least in-part by extracting text from the set of search result items (see, e.g., page 3, lines 4-9; Fig. 4, block 96; and page 10, lines 3-5); and
- using the grammar to interpret an utterance by the user of an additional query term (see, e.g., page 3, lines 4-9; Fig. 4, block 98; and page 10, lines 5-7).

Claim 43

Claim 43 involves a method that enables a user to efficiently conduct a search using a telephone. The user is prompted to depress a sequence of telephone keypad entries corresponding to a sequence of characters of a query term of a search query, and is also prompted to utter the search query. The utterance of the search query is interpreted using a voice recognition grammar that corresponds to the sequence of keys depressed by the user.

More specifically, Claim 43 recites a method for facilitating database searches conducted over a telephone, the method comprising:

- prompting a user to depress a sequence of telephone keypad keys corresponding to a sequence of characters of a query term of a search query, and identifying a resulting sequence of keys depressed by the user (see, e.g., page 2, lines 6-12; page 2, line 24 to page 3, line 3; Fig. 3, block 80; and page 8, line 17 to page 9, line 4);
- prompting the user to utter the search query by voice, and receiving a resulting voice utterance from the user (see, e.g., page 2, line 12; Fig. 1, block 28; page 6, lines 24-26; Fig. 3, block 90; and page 9, lines 16-20); and
- interpreting the voice utterance using a voice recognition grammar that corresponds to the sequence of keys depressed by the user, said voice recognition grammar specifying valid utterances (see, e.g., page 2, lines 13-23; Fig. 1, block 28; page 6, line 10 to page 7, line 10; Fig. 3, block 90; and page 9, lines 18-20).

Claim 50

Claim 50 describes the processing that is performed to interpret a search query submitted by a user by telephone. The user specifies a subset of the characters of the search query, at least in part, using a telephone keypad. The user also utters the search query. The user's utterance of the search query is interpreted using a voice recognition grammar that corresponds to the user's indication of the subset of characters.

More specifically, Claim 50 recites a method of capturing a search query specified by a user by telephone, the method comprising:

- receiving from the user an indication of a subset of the characters contained in the search query, said indication of the subset of characters being specified at least in part as telephone keypad entries (see, e.g., page 2, lines 6-12; page 2, line 24 to page 3, line 3; Fig. 3, block 80; and page 8, line 17 to page 9, line 4);
- receiving from the user a voice utterance that represents the entire search query (see, e.g., page 2, line 12; Fig. 1, block 28; page 6, lines 24-26; Fig. 3, block 90; and page 9, lines 16-20); and

- interpreting the voice utterance using a voice recognition grammar that corresponds to the indication of the subset of characters, said voice recognition grammar specifying valid utterances (see, e.g., page 2, lines 13 and 14; Fig. 1, block 28; page 6, line 10 to page 7, line 10; Fig. 3, block 90; and page 9, lines 18-20).

VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

The following grounds for rejection are to be reviewed on appeal.

1. The rejection of Claims 1-55 under 35 U.S.C. 103(a) over U.S. Patent 6,377,927 ("Loghmani") in view of U.S. Patent 5,917,889 ("Brotman"), and further in view of U.S. Patent 6,434,524 ("Weber '524"). The Examiner also appears to rely on a fourth reference, U.S. Patent 6,532,444 ("Weber '444"), in connection with some of these claims, namely dependent claims 17-19, 20 and 27 (see Current Office Action at page 9).

2. The rejection of Claim 50 under 35 U.S.C. 103(a) over Brotman in view of U.S. Patent 6,839,669 ("Gould").

In connection with these rejections, Appellant requests review of both (1) the Examiner's conclusion that the references teach or suggest all of the claim limitations, and (2) the Examiner's conclusion that a teaching, suggestion, or motivation exists to combine the references.

With respect to each reference that is not prior art under § 102(b), Appellant will treat the reference as prior art for purposes of this appeal, but reserves the right to later disqualify the reference as prior art.

VII. ARGUMENT

In rejecting claims under 35 U.S.C. § 103, the Examiner bears the initial burden of presenting a prima facie case of obviousness. See In re Rijckaert, 9 F.3d 1531, 1532, 28 USPQ2d 1955, 1956 (Fed. Cir. 1993). The Examiner may satisfy this burden only by showing some objective teaching in the prior art or that knowledge generally available to one of ordinary skill in the art. In re Fine, 837 F.2d 1071, 1074, 5 USPQ2d 1596, 1598 (Fed. Cir. 1988).

As set forth below, the Examiner has failed to satisfy this burden with respect to each of the rejected claims. By declining to present arguments with respect to some of the dependent

claims, Appellant does not imply that the limitations added by such claims are taught or suggested by the references.

1. Rejection of Claims 1-55 over Loghmani, Brotman, and Weber '524

As set forth below, the rejections of Claim 1-55 over the combination of Loghmani, Brotman, and Weber '524 are improper because the references do not teach or suggest all of the claim limitations, and also because the Examiner has not identified a teaching, suggestion or motivation to combine Loghmani and Brotman.

a. Loghmani

Loghmani discloses a voice-optimized database system that enables users to conduct database searches by voice. Each searchable item in the database is stored in association with an audio vector that characterizes the sound of a name or phrase associated with the item. The audio vector includes vector components having values for respective phonemes in the searchable item's name or phrase. See col. 4, lines 13-37. Multiple audio vectors may be stored for a given searchable item, each of which corresponds to a different phrase that may be uttered to search for the item.

To process a spoken query from a user, the spoken query is parsed based on the phonemes therein, and an audio vector is assigned to the spoken query. The assigned audio vector is then compared to the audio vectors associated with the searchable items in the database to search for items having an audio vector that is close to that of the search query. See col. 4, lines 38-55. The search results are then presented to the user. Thus, the spoken search query is processed without using a voice recognition grammar to initially convert the search query to text.

Loghmani also discloses that if the database does not support the use of audio vectors, an intermediate audio vector valuation module may be used to convert the phonemes in the spoken query to text, so that one or more textual versions of the query may be passed to the database. See column 8, lines 38-55.

b. Brotman

Brotman discloses techniques for reliably capturing a string of characters specified by a user via a telephone. These techniques involve having the user both (1) utter all of the characters in the intended string, and (2) select the corresponding keys on the telephone keypad for each of these characters—either by depressing these keys or by uttering the number (0-9) of each such

key. The character utterances and the keypad selections are then used in combination to predict the characters intended by the user, so that the likelihood of misrecognition is reduced.

For example, to input the word "cat," the caller would utter the letters C-A-T, and would depress the corresponding telephone keys 2-2-8 (or alternatively, utter the numbers 2-2-8). The keypad selections would then be used to limit the possible interpretations of the character utterances. For example, to interpret the utterance of the letter "C," the system may treat "A," "B" and "C" (corresponding to the key for "2") as the only valid utterances. See column 3, lines 42-53. This is accomplished by using the telephone keypad entries to create a grammar that specifies the valid characters that may be uttered by the user. See column 4, lines 36-41.

Using this process, Brotman's system generates a string of characters that are predicted to be the characters intended by the user. This character string is then audibly output to the user, and the user is prompted to indicate, with a "yes" or "no" reply, whether the generated string is what the user intended. If the string has not been accurately captured, the caller can continue to interact with the system until the intended string has been correctly identified. See Figure 2, blocks 670-730, and column 5, line 36 to column 6, line 10.

Brotman is not directed to the capture of search queries. Even if Brotman's method were used to capture search queries, it would not provide an efficient process for doing so. For example, a user wishing to submit the search query "Stephen King" would apparently have to utter all eleven letters of the query, and would also have to select the corresponding eleven keys on the telephone keypad. In contrast, in Appellant's preferred embodiment, the user could conduct this search by pressing the telephone keys containing the characters "S-T-E" (and/or uttering these characters) and uttering the name "Stephen King." The increase in efficiency over Brotman is even greater for longer queries.

Nothing can be taken from Loghmani that would improve this inefficiency in Brotman's character entry process. To the contrary, as discussed below, the combination of Brotman and Loghmani proposed by the Examiner would actually *increase* the number of steps the user would need to perform to input the desired character string, resulting in a less efficient process.

c. Weber '524

Weber '524 discloses a user interface through which a user can interact with a computer system by voice. The user's utterances are interpreted in-part using context-specific voice

recognition grammars that correspond to specific subjects such as "news," "weather," and "stocks." Col. 7, lines 14-30. These context-specific grammars are not selected based on the user's entry of specific characters of the uttered term or phrase. Rather, they are apparently selected based on the topic or subject currently being browsed, as determined from prior utterances. See col. 3, lines 15-21 and col. 8, lines 45-53.

In contrast to the methods disclosed in the present application, Weber's method of using context-specific voice recognition grammars is not well suited for searching large domains of items, such as a domain of millions of book titles or music titles. If Weber's method were used for this purpose, the user would likely have to "drill down" through multiple levels of item categories and subcategories (e.g., books\fiction\mysteries); otherwise, the voice recognition grammars would most likely be too large to provide reliable voice recognition. In addition to being burdensome to users, such an approach would require the users to know how the items they are searching for are categorized. Brotman and Loghmani do not suggest a solution to this deficiency in Weber '524.

Independent Claim 1

Claim 1 is directed to an embodiment in which the grammar used to interpret a voice query (i.e., an utterance of a search query) is a "dynamic grammar" that is generated after the user has submitted a set of characters that define a portion of the search query. The dynamic grammar is generated based at least in part on an identified subset of items that correspond to the set of characters received from the user. The claim reads as follows, with reference characters added for purposes of discussion:

> 1. A method for improving voice recognition accuracy when a user submits a search query by voice to search a domain of items, the method comprising:
>
> (a) prompting a user to submit a set of characters of a voice query for searching the domain of items, and receiving the set of characters from the user, wherein the voice query is an utterance by the user of a search query, and the set of characters defines a portion of the search query;
>
> (b) in response to receiving the set of characters from the user, identifying a subset of items in the domain that correspond to the set of characters;
>
> (c) generating a dynamic grammar based at least in part on the subset of items, said grammar specifying valid utterances for interpreting the voice query;
>
> (d) prompting the user to submit the voice query, and receiving the voice query from the user; and

(e) interpreting the voice query using the dynamic grammar.

The rejection of Claim 1 is improper because, among other reasons, the applied references (Loghmani, Brotman, and Weber '524) do not teach or suggest prompting a user to submit a set of characters that "defines a portion of the search query" as recited in subparagraph (a). In connection with these limitations, Appellant submits that the term "portion" has its ordinary meaning, which is "a section or quantity within a larger thing; a part of a whole." See dictionary.com, definition no. 1 for "portion."

In connection with this claim language, the Examiner relies on the character-string entry method disclosed in Brotman. Brotman, however, does not disclose a method in which the user is prompted to submit a set of characters "that defines a portion of" a search query or other intended character string. Rather, Brotman describes a character entry method in which the user both utters, and selects the corresponding telephone keys for, all of the characters of the intended character string.

The Examiner responds to Appellant's arguments regarding the term "portion" by pointing out that the claim language recites "a portion of the search query," and not "a portion of the characters" of the search query. Current Office Action at page 13, lines 2-4. Appellant respectfully submits that the two are synonymous. In this regard, Appellant submits that "search query" is properly construed in Claim 1 to refer to the complete textual representation of the voice query.

The Examiner also responds by pointing to an example in Brotman in which the letters "IRA" represent a subset of the 27 possible character sequences corresponding to the digits "4", "7", "2" entered by the user. Current Office Action at page 13. The Examiner does not, however, explain how this example teaches or suggests the claim limitations at issue.

The rejection of Claim 1 is also improper because the applied references do not teach or suggest the limitations of subparagraph (c), namely "generating a dynamic grammar based at least in part on the subset of items, said grammar specifying valid utterances for interpreting the voice query." In connection with these limitations, the Examiner relies on Figure 2, block 630 of Brotman. As explained at column 4, lines 36-41 of Brotman, block 630 of Figure 2 depicts a step in which each numerical telephone digit (0-9) selected by the user is used to limit the

possible characters that can be validly uttered by the user. For instance, if the user depresses the keys "4," "7," "2," a grammar would be created that defines the allowable characters as {G,H,I}, {P,R,S}, and {A,B,C}. With this approach, the grammar is generated based solely on the numerical digits specified by the user. In contrast, Claim 1 recites a method in which the grammar is generated based at least in part on a subset of items, within a domain of items being searched, that correspond to a set of characters received from the user.

Because Claim 1 includes limitations that are not taught or suggested by the applied references, the obviousness rejection of Claim 1 is improper.

The rejection of Claim 1 is also improper because the Examiner has failed to identify a teaching, suggestion or motivation to combine Loghmani and Brotman. See M.P.E.P. § 2143.03. The basis given by the Examiner for combining these two references is that the addition of Brotman's character entry process to Loghmani's voice-based search process would "reduce the domain of field choices, as well as improving the accuracy process in using the dual character input and follow-up speech verification" of Brotman. Current Office Action at pages 4 and 5.

This asserted basis, however, fails to recognize that if Brotman's process were used to capture and verify a search query string intended by a user, there would be no need to additionally use Loghmani's process to interpret the user's utterance of the full search query. This is because Brotman's process allows the user to verify that the intended character string has been properly identified by the system. Once the search query string has been properly identified, it can be processed using conventional textual query processing methods, and Loghmani's spoken query processing methods become unnecessary.

Further, if both the Brotman process and the Loghmani process were used in combination to capture the intended search query, the user would apparently have to go through the unnecessarily burdensome process of (1) selecting the telephone keypad keys associated with all of the characters of the search query string, (2) uttering each character of the search query string, (3) verifying that the search string has been properly recognized by the system, and (4) uttering the search query. One skilled in the art would not be motivated to design a system that requires users to undergo such an unnecessarily burdensome process.

In the Current Office Action, the Examiner responds to Appellant's argument in-part by stating the following: "The character identification of Brotman would further resolve any voice

recognition ambiguities, and the voice recognition aspect of Loghmani would further limit any character entry ambiguities." Current Office Action at page 12, lines 7-9. Nothing in either reference, however, suggests that the processes describe therein suffer from voice recognition ambiguity problems. Thus, even if the Examiner's statement is accurate, it does not identify a motivation to combine. This is particularly true given the reduction in efficiency (increased number of user steps) that would apparently result from the combination.

The Examiner also responds to Appellant's arguments by asserting that Loghmani produces multiple possible results for the query, thereby improving the accuracy and speed of Brotman's system. In addition, the Examiner asserts that "reversing the two references would also lead to an improvement in the recognition accuracy and speed, because the voice recognition would reduce the number of interactions of the user with character recognition results (in the Brotman reference)." Current Office Action at page 12. As explained above, however, neither reference indicates that the processes described therein suffer from voice recognition accuracy problems or associated speed problems. Thus, even if true, these statements do not identify a motivation to combine, particularly given the apparent reduction in efficiency that would result.

Because the Examiner has not identified a teaching, suggestion or motivation to combine Brotman and Loghmani, the obviousness rejection of Claim 1 is improper.

In the Supplemental Appeal Brief, Appellant asserted that it is not clear how the Examiner is relying on Weber '524 in connection with the independent claims. Supplemental Appeal Brief at page 5, third paragraph. In response, the Examiner explains that Weber '524 is introduced for purposes of updating the dynamic grammar. Current Office Action at page 12, lines 2-4. Appellant respectfully submits that the Examiner's explanation does not clarify this issue.

In summary, the Examiner has not established prima facie obviousness with respect to Claim 1, both because the applied references do not teach or suggest all of the claim limitations, and because the Examiner has not identified a teaching, motivation or suggestion to combine Loghmani and Brotman.

Dependent Claims 2-14, 39 and 40

Claims 2-14, 39 and 40 depend directly or indirectly from Claim 1. Thus, the rejections of Claims 2-14, 39 and 40 are improper for the reasons set forth above for Claim 1. The

rejections of these dependent claims are also improper, and should be reversed, for the additional reasons set forth below for specific claims.

Dependent Claim 2

Claim 2 depends from Claim 1 and adds the following: "wherein prompting a user to submit a set of characters comprises prompting the user to submit the first N characters of a query term, where N is greater than 1." In connection with this claim, the Examiner cites Brotman at col. 4, lines 36-41. Brotman, however, teaches prompting the user to submit all of the characters of the string, which is not the same as "prompting the user to submit the first N characters." Thus, the Examiner has not shown that the limitations of Claim 2 are taught or suggested.

Dependent Claim 9

Claim 9 depends from Claim 1 and adds the following: "wherein generating a dynamic grammar comprises extracting text from the subset of items." In connection with this claim, the Examiner cites Brotman at col. 4, lines 47-52. While this section of Brotman may discuss the use of grammars (particularly for recognizing utterances of characters), it says nothing about extracting text from a subset of items. Indeed, the grammars at issue in Brotman do not appear to be generated from extracted text. Thus, the Examiner has not shown that the limitations of Claim 9 are taught or suggested.

Dependent Claim 13

Claim 13 depends from Claim 1 and adds the following: "wherein receiving the set of characters comprises determining in real time whether a number of entered characters is sufficient to produce a grammar that falls below a threshold size." In connection with this claim, the Examiner cites Brotman at col. 5, lines 25-30. Appellant respectfully submits that the cited portion of Brotman has no relevance whatsoever to the limitations of Claim 13, and that the Examiner has not shown that these limitations are taught or suggested. Indeed, Brotman does not teach or suggest the limitations recited in Claim 13.

Dependent Claim 14

Claim 14 depends from Claim 1 and adds the following: "executing a search using the voice query as interpreted using the dynamic grammar to identify a set of search result items; providing the user an option to add an additional query term to the voice query to refine the

search; generating a second dynamic grammar at least in part from the set of search result items; and receiving a voice entry of the additional query term from the user, and interpreting the voice entry using the second dynamic grammar." These limitations are not taught or suggested by the applied references, and the Examiner has not attempted to show otherwise.

Dependent Claim 39

Claim 39 depends from Claim 1, and is patentable over the applied references for the same reasons as set forth for Claim 1. In addition, the rejection of Claim 39 is improper because the applied references do not teach or suggest that "the set of characters is a subset of the characters contained in a textual representation of the voice query." The Examiner cites Brotman at col. 5, lines 25-30, but makes no attempt to show how or why the limitations of Claim 39 are taught or suggested.

Independent Claim 15

Claim 15 is directed to a method for improving voice recognition accuracy when a user submits a query by voice to search a domain of items. The method comprises "receiving a set of characters entered by a user, the set of characters representing a portion of a query." The applied references (Loghmani, Brotman, and Weber '524) do not teach or suggest this step. As mentioned above, if Brotman's character entry method were used to capture a query, the received characters would be those of the entire query, and not "a portion" of the query.

Claim 15 also recites "in response to receiving the set of characters, selecting a grammar which is derived at least in-part from text extracted from a subset of items that correspond to the set of characters entered by the user;" and "providing the grammar to a voice recognition system for use in interpreting the query as entered by the user by voice." The applied references do not teach or suggest these limitations. The Examiner did not fully address these limitations in the Current Office Action.

The applied references also fail to teach or suggest a method in which "the user's entry of a subset of characters of the query, together with the user's utterance of the full query, are used in combination to capture the query." In connection with the term "subset," the Examiner points to an example in Brotman in which the letters "IRA" represent a subset of the 27 possible letter sequences corresponding to the digits "4", "7", "2" entered by the user. Current Office Action at

page 13. The Examiner does not, however, explain how this example teaches or suggests the claim limitations at issue.

Because Claim 15 includes limitations that are not taught or suggested by the applied references, the obviousness rejection of Claim 15 is improper.

The obviousness rejection of Claim 15 is also improper because, as explained above in connection with Claim 1, the Examiner has not identified a teaching, suggestion or motivation to combine Loghmani with Brotman.

In summary, the Examiner has not established prima facie obviousness with respect to Claim 15, both because the applied references do not teach or suggest all of the claim limitations, and because the Examiner has not identified a teaching, motivation or suggestion to combine Loghmani and Brotman.

Dependent Claims 16-23 and 41

Claims 16-23 and 41 depend directly or indirectly from Claim 15. Thus, the rejections of Claims 16-23 and 41 are improper for the reasons set forth above for Claim 15. The rejections of these dependent claims are also improper, and should be reversed, for the additional reasons set forth below for specific claims.

Dependent Claim 16

Claim 16 depends from Claim 15, and adds the following: "wherein selecting a grammar comprises: executing an initial search to identify the subset of items that correspond to the set of characters; and extracting text from the subset of items for incorporation into the grammar." These limitations are not taught or suggested by the applied references, and are not fully addressed in the Current Office Action.

Dependent Claim 23

Claim 23 depends from Claim 15, and adds the following: "wherein receiving a set of characters comprises determining in real time whether a number of entered characters is sufficient to produce a grammar that falls below a threshold size." In connection with this claim, the Examiner cites Brotman at col. 5, lines 25-30. Appellant respectfully submits that the cited portion of Brotman has no relevance whatsoever to the limitations of Claim 23, and that Brotman does not teach or suggest these limitations.

Independent Claim 24

Claim 24 is directed to a system that includes "a first code module which causes a user to be prompted to enter a set of characters of a query such that the user may partially specify the query," and "a second code module which causes the user to be prompted to utter the query." The claim also calls for a query server that "is programmed to use the set of characters to select a grammar for use by [a] voice recognition system to interpret the query as uttered by the user."

The applied references do not teach or suggest such a system. In this regard, if Brotman's character entry method were used to capture search queries, the user would not enter a set of characters that partially specifies the query, but rather would individually specify all of the characters of the search query.

Further, even if Brotman's method were used to capture a set of characters of the query, it would not be obvious from the applied references to then use this captured set of characters to select a grammar for interpreting the user's utterance of the query. The Examiner has not identified any disclosure in the applied references that suggests this aspect of the claimed method.

Because Claim 24 includes limitations that are not taught or suggested by the applied references, the obviousness rejection of Claim 24 is improper.

The obviousness rejection of Claim 24 is also improper because, as explained above in connection with Claim 1, the Examiner has not identified a teaching, suggestion or motivation to combine Loghmani with Brotman.

In summary, Appellant respectfully submits that the Examiner has not established prima facie obviousness with respect to Claim 24, both because the applied references do not teach or suggest all of the claim limitations, and because the Examiner has not identified a teaching, motivation or suggestion to combine Loghmani and Brotman.

Dependent Claims 25-32

Claims 25-32 depend directly or indirectly from Claim 24. Thus, the rejections of Claims 25-32 are improper for the reasons set forth above for Claim 24. The rejections of these dependent claims are also improper, and should be reversed, for the additional reasons set forth below for specific claims.

Dependent Claim 25

Claim 25 depends from Claim 24, and adds the following: "wherein the first and second code modules comprise voiceXML coding." These limitations are not taught or suggested by the applied references, and are not fully addressed in the Current Office Action.

Dependent Claim 26

Claim 26 depends from Claim 24, and adds the following: "wherein the query server selects the grammar by at least: executing a preliminary search to identify a subset of items that match the set of characters; and extracting text from the subset of items to incorporate into grammar." These limitations are not taught or suggested by the applied references, and are not fully addressed in the Current Office Action.

Dependent Claim 27

Claim 27 depends from Claim 24, and adds the following: "wherein the query server is programmed to extract author names from the subset of items to generate a grammar for performing a voice-based author search." In connection with these limitations, the Examiner asserts that both Weber '444 and Loghmani teach the use of subcategories labeled "author." (As mentioned above, the Examiner appears to rely informally on Weber '444 in connection with this and several other dependent claims.) Even if true, however, this assertion does not show how or why it would be obvious "to extract author names from the subset of items to generate a grammar for performing a voice-based author search" as claimed. Thus, the Examiner has not shown that the limitations of Claim 27 are taught or suggested. Indeed, Brotman, Loghmani, Weber '524 and Weber '444 do not teach or suggest the limitations of Claim 27.

Dependent Claim 29

Claim 29 depends from Claim 24, and adds the following: "wherein the set of characters is a set of the first N letters of a query term, where N is greater than 1." In connection with this claim, the Examiner cites Brotman at col. 4, lines 36-41. Brotman, however, teaches prompting the user to submit all of the characters of the string, which is not the same as, or suggestive of prompting the user to enter the first N letters." Thus, the Examiner has not shown that the limitations of Claim 29 are taught or suggested. Indeed, Brotman does not teach or suggest these limitations.

Dependent Claim 31

Claim 31 depends from Claim 29, and adds the following: "wherein N is selected based on a target grammar size." In connection with this claim, the Examiner cites Brotman at col. 5, lines 25-30. Appellant respectfully submits that the cited portion of Brotman has no relevance whatsoever to the limitations of Claim 31, and that Brotman does not teach or suggest these limitations.

Independent Claim 33

Claim 33 involves a method in which a user refines a search query by uttering an additional query term to add to the query. To interpret the user's utterance of the additional query term, a grammar is generated, at least in-part, by extracting text from the set of search result items resulting from the query. A preferred embodiment of this method is described beginning at page 9, line 24 of the present application, and is shown primarily in Figure 4.

The rejection of Claim 33 is improper because, among other reasons, the applied references do not teach or suggest the following limitations: "providing the user an option to refine the query by adding an additional query term," "generating a grammar at least in-part by extracting text from the set of search result items," and "using the grammar to interpret an utterance by the user of an additional query term." The Current Office Action does not fully address these limitations. Indeed, none of the applied references discloses a query refinement process, let alone the particular query refinement process defined in Claim 33.

The obviousness rejection of Claim 33 is also improper because, as explained above in connection with Claim 1, the Examiner has not identified a teaching, suggestion or motivation to combine Loghmani with Brotman.

Dependent Claims 34-38 and 42

Claims 34-38 and 42 depend directly or indirectly from Claim 33. Thus, the rejections of Claims 34-38 and 42 are improper for the reasons set forth above for Claim 33.

Independent Claim 43

Claim 43 is directed to a method that involves "prompting a user to depress a sequence of telephone keypad keys corresponding to a sequence of characters of a query term of a search query." The user is also prompted "to utter the search query by voice." The voice utterance of

the search query is interpreted "using a voice recognition grammar that corresponds to the sequence of keys depressed by the user."

Of the applied references, Brotman is the only reference that involves the use of a voice recognition grammar that is based on telephone keypad keys depressed by a user. Brotman's grammar, however, is suitable only for interpreting utterances of individual characters, and not for interpreting an utterance of a search query. Thus, even if Brotman's teaching to create a grammar based on a user's telephone keypad entries were combined with the teachings of Loghmani and/or Weber '524, the combination would not involve "interpreting the voice utterance [of the search query] using a voice recognition grammar that corresponds to the sequence of keys depressed by the user," as required by Claim 43.

Because Claim 43 includes limitations that are not taught or suggested by the applied references, the obviousness rejection of Claim 43 is improper.

The obviousness rejection of Claim 43 is also improper because, as explained above in connection with Claim 1, the Examiner has not identified a teaching, suggestion or motivation to combine Loghmani with Brotman.

Dependent Claims 44-49

Claims 44-49 depend directly or indirectly from Claim 43. Thus, the rejections of Claims 44-49 are improper for the reasons set forth above for Claim 43. The rejections of these dependent claims are also improper, and should be reversed, for the additional reasons set forth below for specific claims.

Dependent Claim 47

Claim 47 depends from Claim 43, and adds the following: "further comprising selecting the voice recognition grammar from a repository of previously-generated voice recognition grammars in which different voice recognition grammars correspond to different sequences of characters." In connection with this claim, the Examiner points to Brotman at col. 4, lines 47-52. This section of Brotman discusses the use of forward and backward grammar prediction, which involves using surrounding characters to predict the character intended by the user. While this portion of Brotman may involve a type of voice recognition grammar, it does not teach or suggest the limitations at issue. Thus, the Examiner has not shown that the applied references teach or suggest the limitations of Claim 47.

Independent Claim 50

Claim 50 is directed to a method of capturing a search query specified by a user by telephone. The method comprises "receiving from the user an indication of a subset of the characters contained in the search query, said indication of the subset of characters being specified at least in part as telephone keypad entries." The method further comprises "receiving from the user a voice utterance that represents the entire search query," and "interpreting the voice utterance using a voice recognition grammar that corresponds to the indication of the subset of characters."

The rejection of Claim 50 is improper because, among other reasons, the applied references do not teach or suggest "interpreting the voice utterance [that represents the entire search query] using a voice recognition grammar that corresponds to the indication of the subset of characters [contained in the search query]." As discussed in connection with Claim 43, Brotman is the only applied reference that involves the use of a voice recognition grammar that is based on telephone keypad entries, and this grammar would not be suitable for interpreting a voice utterance of a search query. Further, Brotman's grammar is created based on the keypad entries corresponding to the entire character string to be captured, and not an indication of a "subset of characters" as claimed.

Because Claim 50 includes limitations that are not taught or suggested by the applied references, the obviousness rejection of Claims 50 is improper.

The obviousness rejection of Claim 50 is also improper because, as explained above in connection with Claim 1, the Examiner has not identified a teaching, suggestion or motivation to combine Loghmani with Brotman.

Dependent Claims 51-55

Claims 51-55 depend directly or indirectly from Claim 50. Thus, the rejections of Claims 51-55 are improper for the reasons set forth above for Claim 50. The rejections of these dependent claims are also improper, and should be reversed, for the additional reasons set forth below for specific claims.

Dependent Claim 52

Claim 52 depends from Claim 50, and adds the following: "further comprising selecting the voice recognition grammar from a repository of previously-generated voice recognition

grammars in which different voice recognition grammars correspond to different sets of characters." The applied references do not teach or suggest these limitations. In connection with this claim, the Examiner points to Brotman at col. 4, lines 15-35. The cited portion of Brotman, however, does not teach the selection of a voice recognition grammar from a repository of previously-generated voice recognition grammars in which different voice recognition grammars correspond to different sets of characters.

Dependent Claim 54

Claim 54 depends from Claim 50, and adds the following: "further comprising executing a database search using a textual representation of the voice utterance." In connection with this claim, the Examiner points to Brotman at col. 5, lines 25-30, but makes no attempt to explain how the cited portion of Brotman discloses the subject limitations.

2. Rejection of Claim 50 over Brotman in view of Gould

Claim 50 also stands rejected on obviousness grounds over Brotman in view of Gould. The claim reads as follows:

> 50. A method of capturing a search query specified by a user by telephone, the method comprising:
>
> receiving from the user an indication of a subset of the characters contained in the search query, said indication of the subset of characters being specified at least in part as telephone keypad entries;
>
> receiving from the user a voice utterance that represents the entire search query; and
>
> interpreting the voice utterance using a voice recognition grammar that corresponds to the indication of the subset of characters, said voice recognition grammar specifying valid utterances.

Appellant respectfully submits that this rejection is improper because, among other reasons, Brotman and Gould do not teach or suggest interpreting the voice utterance "using a voice recognition grammar that corresponds to the indication of the subset of characters." In connection with these limitations, and particularly the phrase "subset of characters," the Examiner points to Figure 34B of Gould. This figure illustrates a scenario in which the user partially spells the misrecognized word "kibitzers" by uttering or typing in each of the first four letters of the word. The program then responds by displaying a list of the words that start with these four letters, and the user then selects the desired word from the list. See Gould at col. 79,

lines 30-56. While the illustrated scenario may involve a subset of the characters of a desired word, it does not involve the use of "a voice recognition grammar that corresponds to the indication of the subset of characters."

In addition, the Examiner does not explain how the addition of this "auto-completion" feature of Gould to Brotman would result in the use of "a voice recognition grammar that corresponds to the indication of the subset of characters." To the contrary, the Examiner appears to suggest that the combination would produce a system in which the user enters the first few characters of the desired string using Brotman's method, and is then presented (presumably audibly by telephone) with a list of the matching words. Current Office Action at page 3, last paragraph. If the references were combined in such a manner, however, it is not clear how or why there would be any need for "a voice recognition grammar that corresponds to the indication of the subset of characters" as claimed.

The rejection is also improper because Brotman and Gould do not collectively teach or suggest either "receiving from the user an indication of... characters contained *in the search query,*" or "receiving from the user a voice utterance that represents *the entire search query*" (emphasis added). In connection with these limitations, and particularly the "search query" limitation, the Examiner points to Brotman at col. 4, lines 5-15 and Fig. 2, block 650. Neither the cited nor any other portion of Brotman, however, discloses or suggests a search query. Thus, the Examiner has not shown that the subject limitations are taught or suggested.

For these reasons, Appellant respectfully submits that the Examiner has not established a prima facie case of obviousness with respect this additional basis for the rejection of Claim 50.

VIII. CONCLUSION

For the reasons set forth above, the rejections of Claims 1-55 are improper and should be reversed.

Respectfully submitted,

KNOBBE, MARTENS, OLSON & BEAR, LLP

Dated: 6-29-06

By: _____

Ronald J. Schoenbaum
Reg. No. 38,297
2040 Main Street, 14th Floor
Irvine, CA, 92614
Customer No. 20,995
949-721-2950

APPENDIX – PENDING CLAIMS

1. A method for improving voice recognition accuracy when a user submits a search query by voice to search a domain of items, the method comprising:

prompting a user to submit a set of characters of a voice query for searching the domain of items, and receiving the set of characters from the user, wherein the voice query is an utterance by the user of a search query, and the set of characters defines a portion of the search query;

in response to receiving the set of characters from the user, identifying a subset of items in the domain that correspond to the set of characters;

generating a dynamic grammar based at least in part on the subset of items, said grammar specifying valid utterances for interpreting the voice query;

prompting the user to submit the voice query, and receiving the voice query from the user; and

interpreting the voice query using the dynamic grammar.

2. The method as in Claim 1, wherein prompting a user to submit a set of characters comprises prompting the user to submit the first N characters of a query term, where N is greater than 1.

3. The method as defined in Claim 1, wherein prompting a user to submit a set of characters comprises prompting the user to submit a set of characters of an author's name.

4. The method as defined in Claim 3, wherein generating a dynamic grammar comprises incorporating into the grammar names of authors of the items within the subset of items.

5. The method as defined in Claim 4, wherein the dynamic grammar consists essentially of the names of the authors of the items within the subset of items.

6. The method as defined in Claim 4, further comprising incorporating into the dynamic grammar non-author terms extracted from the subset of items.

7. The method as defined in Claim 1, wherein prompting a user to submit a set of characters comprises prompting the user to select the characters on a telephone keypad.

8. The method as defined in Claim 7, wherein prompting a user to submit a set of characters further comprises prompting the user to utter the characters, and wherein receiving the set of characters comprises using the keypad entries of the characters to interpret utterances by the user of the characters.

9. The method as defined in Claim 1, wherein generating a dynamic grammar comprises extracting text from the subset of items.

10. The method as defined in Claim 9, wherein extracting text from the subset of items comprises extracting the text from a database field corresponding to a search context of the query.

11. The method as defined in Claim 1, further comprising storing the dynamic grammar within a cache for subsequent use.

12. The method as defined in Claim 1, wherein prompting a user to submit a set of characters comprises prompting the user to enter a fixed number of characters, wherein the fixed number is selected based on a target grammar size.

13. The method as defined in Claim 1, wherein receiving the set of characters comprises determining in real time whether a number of entered characters is sufficient to produce a grammar that falls below a threshold size.

14. The method as defined in Claim 1, further comprising:

executing a search using the voice query as interpreted using the dynamic grammar to identify a set of search result items;

providing the user an option to add an additional query term to the voice query to refine the search;

generating a second dynamic grammar at least in part from the set of search result items; and

receiving a voice entry of the additional query term from the user, and interpreting the voice entry using the second dynamic grammar.

15. A method for improving voice recognition accuracy when a user submits a query by voice to search a domain of items, the method comprising:

receiving a set of characters entered by a user, the set of characters representing a portion of a query;

in response to receiving the set of characters, selecting a grammar which is derived at least in-part from text extracted from a subset of items that correspond to the set of characters entered by the user; and

providing the grammar to a voice recognition system for use in interpreting the query as entered by the user by voice;

whereby the user's entry of a subset of characters of the query, together with the user's utterance of the full query, are used in combination to capture the query.

16. The method as defined in Claim 15, wherein selecting a grammar comprises:

executing an initial search to identify the subset of items that correspond to the set of characters; and

extracting text from the subset of items for incorporation into the grammar.

17. The method as defined in Claim 16, wherein extracting text from the subset of items comprises extracting the text from a database field corresponding to a search context of the query.

18. The method as defined in Claim 17, wherein the search context comprises an author search, and the database field is an author field.

19. The method as defined in Claim 15, wherein selecting a grammar comprises reading a previously generated grammar from memory based on the set of characters entered by the user.

20. The method as in Claim 15, wherein receiving a set of characters comprises receiving the first N characters of a query term, where N is greater than 1.

21. The method as in Claim 15, wherein receiving a set of characters comprises receiving characters entered at least in-part using a telephone keypad.

22. The method as in Claim 15, wherein receiving a set of characters comprises using a telephone keypad entry of a character by the user to interpret an utterance of the character by the user.

23. The method as defined in Claim 15, wherein receiving a set of characters comprises determining in real time whether a number of entered characters is sufficient to produce a grammar that falls below a threshold size.

24. A system for conducting searches by voice, comprising:

a database of items;

a query server which searches the database of items according to voice queries from users, the query server coupled to a voice recognition system which interprets the voice queries according to grammars;

a first code module which causes a user to be prompted to enter a set of characters of a query such that the user may partially specify the query; and

a second code module which causes the user to be prompted to utter the query;

wherein the query server is programmed to use the set of characters to select a grammar for use by the voice recognition system to interpret the query as uttered by the user.

25. The system as defined in Claim 24, wherein the first and second code modules comprise voiceXML coding.

26. The system as defined in Claim 24, wherein the query server selects the grammar by at least:

executing a preliminary search to identify a subset of items that match the set of characters; and

extracting text from the subset of items to incorporate into grammar.

27. The system as defined in Claim 26, wherein the query server is programmed to extract author names from the subset of items to generate a grammar for performing a voice-based author search.

28. The system as defined in Claim 24, wherein the query server is programmed to select the grammar from memory using the set of characters.

29. The system as defined in Claim 24, wherein the set of characters is a set of the first N letters of a query term, where N is greater than 1.

30. The system as defined in Claim 29, wherein the query term is a name of an author.

31. The system as defined in Claim 29, wherein N is selected based on a target grammar size.

32. The system as defined in Claim 24, wherein the first code module prompts the user to both utter, and enter on a telephone keypad, each alphabetic character of the set.

33. A method of assisting users in locating items in a database using voice queries, the method comprising:

receiving a voice query from a user, and identifying a set of search result items that are responsive to the voice query;

providing the user an option to refine the query by adding an additional query term;

generating a grammar at least in-part by extracting text from the set of search result items; and

using the grammar to interpret an utterance by the user of an additional query term.

34. The method as defined in Claim 33, wherein generating a grammar comprises extracting text from a database field corresponding to a search context of the query.

35. The method as defined in Claim 33, wherein using the grammar to interpret an utterance comprises using the grammar to interpret utterances of multiple additional query terms by the user.

36. The method as defined in Claim 33, wherein the grammar is generated in response to selection by the user of the option to add an additional query term.

37. The method as defined in Claim 33, wherein the option to refine the query is presented to the user only if the number of items in the set exceeds a predefined threshold.

38. The method as defined in Claim 33, further comprising storing the grammar in a cache for use with subsequent query submissions.

39. The method as in Claim 1, wherein the set of characters is a subset of the characters contained in a textual representation of the voice query.

40. A system that operates according to the method of Claim 1.

41. A system that operates according to the method of Claim 15.

42. A system that operates according to the method of Claim 33.

43. A method for facilitating database searches conducted over a telephone, the method comprising:

prompting a user to depress a sequence of telephone keypad keys corresponding to a sequence of characters of a query term of a search query, and identifying a resulting sequence of keys depressed by the user;

prompting the user to utter the search query by voice, and receiving a resulting voice utterance from the user; and

interpreting the voice utterance using a voice recognition grammar that corresponds to the sequence of keys depressed by the user, said voice recognition grammar specifying valid utterances.

44. The method of Claim 43, wherein the search query consists of said query term.

45. The method of Claim 43, wherein the search query contains multiple query terms.

46. The method of Claim 43, further comprising prompting the user to utter said sequence of characters by voice, and using resulting voice utterances of the characters in combination with the sequence of keys depressed by the user to identify the sequence of characters intended by the user.

47. The method of Claim 43, further comprising selecting the voice recognition grammar from a repository of previously-generated voice recognition grammars in which different voice recognition grammars correspond to different sequences of characters.

48. The method of Claim 43, further comprising generating the voice recognition grammar on-the-fly based on input from the user.

49. A system that operates according to the method of Claim 43.

50. A method of capturing a search query specified by a user by telephone, the method comprising:

receiving from the user an indication of a subset of the characters contained in the search query, said indication of the subset of characters being specified at least in part as telephone keypad entries;

receiving from the user a voice utterance that represents the entire search query; and

interpreting the voice utterance using a voice recognition grammar that corresponds to the indication of the subset of characters, said voice recognition grammar specifying valid utterances.

51. The method of Claim 50, wherein the indication of the subset of characters further comprises respective voice utterances of the characters in the subset.

52. The method of Claim 50, further comprising selecting the voice recognition grammar from a repository of previously-generated voice recognition grammars in which different voice recognition grammars correspond to different sets of characters.

53. The method of Claim 50, further comprising generating the voice recognition grammar on-the-fly in response to input from the user.

54. The method of Claim 50, further comprising executing a database search using a textual representation of the voice utterance.

55. A system that operates according to the method of Claim 50.

EVIDENCE APPENDIX

None

RELATED PROCEEDINGS APPENDIX

None